

Illusory Conjunctions in the Time Domain and the Resulting Time-Course of the Attentional Blink

Juan Botella, Isabel Arend, and Manuel Suero
Autonomous University of Madrid

Illusory conjunctions in the time domain are errors made in binding stimulus features presented in the same spatial position in Rapid Serial Visual Presentation (RSVP) conditions. Botella, Barriopedro, and Suero (2001) devised a model to explain how the distribution of responses originating from stimuli around the target in the series is generated. They proposed two routes consisting of two sequential attempts to make a response. The second attempt (sophisticated guessing) is only employed if the first one (focal attention) fails in producing an integrated perception. This general outline enables specific predictions to be made and tested related to the efficiency of focal attention in generating responses in the first attempt. Participants had to report the single letter in an RSVP stream of letters that was presented in a previously specified color (first target, T1) and then report whether an X (second target, T2) was or was not presented. Performance on T2 showed the typical U-shaped function across the T1-T2 lag that reflects the attentional blink phenomenon. However, as was predicted by Botella, Barriopedro, and Suero's model, the time-course of the interference was shorter for trials with a correct response to T1 than for trials with a T1 error. Furthermore, longer time-courses of interference associated with pre-target and post-target errors to the first target were indistinguishable.

Keywords: illusory conjunctions, features binding, attention

Se llaman conjunciones ilusorias en el dominio del tiempo a los errores que se producen al combinar rasgos estímulares presentados en la misma posición espacial en condiciones de Presentación Rápida de Series Visuales (PRSV). Botella, Barriopedro y Suero (2001) han formalizado un modelo para explicar cómo se genera la distribución de las respuestas de acuerdo con su posición de origen respecto a la posición del blanco o diana en la serie. Propusieron dos rutas que están constituidas por dos intentos secuenciales de alcanzar una respuesta. El segundo intento (una forma de adivinación sofisticada) sólo se utiliza si fracasa el primero (basado en la atención focal) en su intento por producir un percepto integrado. Esta estructura general permite derivar y poner a prueba predicciones concretas relativas al rendimiento de la atención focal en la generación de respuestas en el primer intento. Los participantes tienen que informar de cuál es la única letra de una serie de letras mostrada en PRSV que apareció en un color previamente especificado (primer blanco, T1) y después indicar si también se presentó una X (segundo blanco, T2). El rendimiento con T2 a través de los diferentes desfases T1-T2 mostró la típica función en forma de U que caracteriza el efecto de attentional blink (AB). Sin embargo, tal y como se predice desde el modelo de Botella, Barriopedro y Suero, el curso temporal de la interferencia fue más corto para los ensayos con una respuesta correcta a T1 que en los ensayos con errores. Además, esos cursos temporales de la interferencia más largos, asociados a los errores pre-blanco y post-blanco al primer blanco, fueron indistinguibles.

Palabras clave: Conjunciones ilusorias, pegado de rasgos, atención

This research was supported by the Ministerio de Ciencia y Tecnología of Spain, project BSO2000-0112. We would like to thank the two reviewers for their helpful comments.

Correspondence concerning this article should be addressed to Juan Botella. Facultad de Psicología, Universidad Autónoma de Madrid, Campus de Cantoblanco, 28049 Madrid, Spain. E-mail: juan.botella@uam.es

The term *illusory conjunctions in the time domain* has been proposed to describe the types of errors made in binding stimulus features presented sequentially in the same spatial location by means of the Rapid Serial Visual Presentation (RSVP) technique (Botella, Barriopedro, & Suero, 2001). In this procedure, participants must report the feature from the to-be-reported dimension (response dimension) of the only stimulus (the target) that has a specific feature in the target-defining dimension (key dimension). Errors are called pre-target and post-target intrusions according to whether the feature which wasn't combined correctly with the target-defining feature was a result of an item being presented before or after the target. The distribution of intrusions for a given participant, or group of participants, in a given experimental condition is labeled as a pre-target or post-target pattern, according to whether there is a predominance of intrusions from pre- or post-target positions, or as a symmetrical pattern if there is no predominance of either type (Botella et al., 2001). Most research regarding this phenomenon is directed towards explaining the behavior that determines the distribution of intrusions around the target and why the empirical distribution changes for different combinations of key and response dimensions.

Botella et al. (2001) proposed a two-stage model to account for patterns of illusory conjunctions over time, in which responses are based on two consecutive attempts to retrieve information from an RSVP stream. Two independent but sequentially applied mechanisms are responsible for both attempts. Although most components of the model have received strong empirical support, there is still a lack of empirical evidence for one of these components, namely, the general architecture involved in the two sequential response processes. In the present paper, we present new evidence to support this architecture. First, we outline the general architecture of the model, but include only those aspects necessary to follow the rationale of the present research (for a more detailed description, see Botella et al., 2001). Second, we explain how the general architecture of the model involves a specific prediction related to the detection accuracy of a second target. Then we describe the experiment and discuss the results.

According to the model, two pre-attentive modules routinely, and in parallel, extract the relevant dimensions of all stimuli in the series; Module K searches for the key target-defining dimension, and Module R records the response features. When Module K detects the target-defining feature, a third mechanism, focal attention, is triggered in order to "capture" the target and integrate its features into a single object or integrated perception. This is the first attempt to generate a response, and a response is generated if focal attention is successful. Given the short SOA (span of attention) in most RSVP tasks, the focusing process is assumed to be frequently interrupted by the presentation of the items which follow in the series. However, the time

taken to develop an integrated perception by means of focal attention can be characterized as a random variable. If the integration process is completed rapidly, before the items which follow interrupt the process, it produces a correct feature combination. However, on trials in which focal attention is not able to produce a response before it is interrupted by the items which follow in the series, a second attempt is made to trigger a response. Two main differences separate both attempts. The first difference is that, whereas the attempt based on successful focal attention is made on-line, the second attempt can be considered an off-line process, although the temporal delay is very short. The second difference is that the second attempt is generated by a fourth mechanism that takes into account the levels of activation of the representations of the response features from the items around the target and the target itself; it employs that information to form a "sophisticated guess" based on the application of Luce's rule to those levels of activation. For the purpose of the present research, it is not important how the sophisticated guessing mechanism works. The only important points are that this mechanism is triggered only after focal attention has failed, and that it is the process that can produce the erroneous combinations we call illusory conjunctions.

An implication of the model, and its proposed general architecture, is that trials in which the response is generated in the first attempt (focal attention is successfully completed) always produce correct responses. Trials in which the response is generated by the second attempt can produce either correct responses or intrusions. As a consequence, any empirical distribution of intrusions can only partially be assigned to the two response routes. Whereas all intrusions are produced by the sophisticated guessing mechanism (second attempt), correct responses are a mixture of first and second attempt responses.

The main focus of the present research is to test a prediction derived from the general architecture of Botella et al.'s (2001) model. We used a combination of two effects discovered by way of modern temporal methods for studying attention (Shapiro, 2001). From our point of view, a main difference between the two ways in which a response is produced is the efficiency of the focal attention process ("efficiency" here refers to the speed at which the process is completed). Let's suppose that we can order trials according to the time invested in focusing attention. If we divide this distribution arbitrarily in half, the lower part of the distribution should include trials with shorter values, and those trials are more likely to be solved in the first attempt. In the upper part of the distribution, we have the trials with longer values that should expose target items to more interference from the items which follow in the series. They are not likely to be identified in the first attempt, so responses in those trials more likely result from the sophisticated guessing mechanism on the second attempt. This covariation makes the classification of trials

into correct responses and errors a credible reflection of classification into two categories: fast and slow focusing of attention.

The rationale of the experiment described here is based on the assumption that the model is a good description of how responses are produced in identifying feature-conjunction targets in an RSVP stream. The differential efficiency of the process of focusing associated with correct responses and errors should show up if an appropriate experimental task is designed. We have attempted to do this by employing the Attentional Blink (AB) paradigm, in which the identification of a first target (T1) impairs performance in detecting a second target (T2) presented in positions temporally close to it (Raymond, Shapiro, & Arnell, 1992; Shapiro, Raymond, & Arnell, 1994). It is assumed that the AB reflects the interference that focal attention devoted to T1 produces on processing T2 (Isaak, Shapiro, & Martin, 1999; Shapiro, Arnell, & Raymond, 1997) and that the AB deficit is a graded function of T1 processing difficulty (Chun & Potter, 1995; Seiffert & DiLollo, 1997). In fact, the time-course of the attentional blink reflects the extended time-course of T1's demands on focal attention (Chun & Potter, 1995; Ward, Duncan, & Shapiro, 1996).

Our procedure consists of adding a second target (the letter X), to the target search task in an RSVP stream which the participant must detect after giving the response to the first target ("identify the letter in a given color and then report whether an X has been presented after the first target"). Our prediction (derived from Botella, Barriopedro & Suero's model, 2001) is that, if we classify responses to the first target as correct responses or intrusions we are, in fact, classifying them also into trials that have associated fast and slow focusing processes, respectively. As a consequence, performance in detecting the second target should reflect this fact. Specifically, if the shape of the AB function reflects the time-course of the processes associated with the application of Focal Attention to the first target, then correct responses should be produced by a faster attentional time-course than errors. This prediction agrees with experimental results obtained with several variants of the AB paradigm (Jolicoeur & Dell'Acqua, 1998; Ward, Duncan, & Shapiro, 1996).

An associated, but not-so-obvious prediction, is that intrusions coming from pre- and post-target positions should show the same time-course, as they are all produced in the second response-generating attempt via the sophisticated guessing mechanism. This prediction is different from what would be expected by other views. More traditional explanations of conjunction errors (e.g., Lawrence, 1971; McLean, Broadbent, & Broadbent, 1983) establish that errors are produced because focal attention is applied at the wrong moment (too early for pre-target errors, too late for post-target errors). However, if this is true, the time course of the mechanism generating errors should be different

according to the type of errors. Pre-target and post-target intrusions should show, respectively, shorter and longer ABs than correct responses, as earlier and later attentional focusing, in an attempt to identify T1, should free attentional focus for T2 earlier and later, respectively.

Notice that we use the AB phenomenon as a tool to study how illusory conjunctions are produced. This is the same logic (but in the opposite order) that Chun (1997) used when he plotted the distribution of errors in perceiving the second target as a function of the time course of the AB (T1-T2 lag). He used the pattern of illusory conjunctions to study how the AB is produced. We used the AB to study how illusory conjunctions are produced.

The correct/error dichotomy in the first response is not a pure classification for the efficiency of focusing in trials (correct responses are composed of fast and slow trials, whereas intrusions are all slow trials). However, if this methodological flaw has any effect when trying to test predictions associated to the efficiency of focusing, it must consist in obscuring the difference. If even in these more stringent circumstances an effect is still observed, a strong argument could be made that the classification can be used to separate successful and unsuccessful attentional focusing.

Method

Participants

Eighteen staff members and graduate students participated as volunteers. All of them had normal or corrected-to-normal vision.

Apparatus

The experiment was run on an IBM compatible PC. The experimental program was written with the MEL program (Schneider, 1988).

Stimuli and materials

Four hundred and twenty series of 16 letters were constructed. We will call the "critical set" the five items composed of the target stimulus plus the two items before and after it. The critical set occupied positions 4-8, 5-9 or 6-10. The letters were presented in five different colors (red, green, white, yellow, and blue), and the background of the screen remained gray throughout the experiment. All five colors were used for the five stimuli in the critical set, assigned randomly for each series. For the rest of the series, the four colors not used for the target in that particular series were employed, randomly, with the only restriction being that there were never two consecutive stimuli with the same color. Positions 6 to 8 were selected for the target in order

to separate the critical set from the ends by several items. This way, the data were not contaminated by the well-known trend for participants to report the first and last items from the series (Lawrence, 1971). Nevertheless, although the participants were able to use the serial position as a cue to select the response item, it cannot explain any eventual difference in the comparisons to be made. The first target (T1) was a letter that was defined by a color cued at the beginning of each trial. The second target (T2) was the letter X that could vary from position +1 to +7 after the first target. To report the first target, participants received a menu with six alternatives including the response features from the target, the two items before and after the target, and the option "don't know." The items were presented in the menu in a random order (the use of the menu makes the experimental paradigm more efficient, avoiding the tendency for participants to be too conservative by selecting the "don't know" alternative frequently; see Botella et al, 2001; Botella & Eriksen, 1992). To report if the X was present, participants were required to press the number 1 on the keyboard for "yes" and the number 0 for "no." There were 48 trials for each T1-T2 condition and 84 (20%) for the no-X condition. In those trials in which the letter X was presented in positions +1 or +2, it was replaced, in the menu, by the letter from position +3, as participants were aware that T1 will never be an X.

Procedure

Participants sat with their eyes about 40 cm from the screen. At the beginning of each trial a horizontal string of three "&" signs appeared at the center of the screen, in the position where the letters would appear and in the target color for that trial. Participants began the series of letters by pressing the space bar. Each colored letter remained on the screen for 83 ms, then was immediately replaced by the following one, and so on, until the end of the trial. Upon completion of the series, a response menu for the first target appeared, containing the letters that pertained to the critical set, but presented in a random order along with the option "don't know" (except for +1 and +2 trials; see above). Participants were required to report which letter appeared in the color specified at the beginning of the trial. After giving their response, the second target question appeared: "Was an X present?" Participants were asked to press the number 1 or 0 on the keyboard for yes or no answers, respectively.

Table 1
Mean Percentages of the Three Types of Responses to T1

Pre	Correct	Post
15.5	61.5	23.0

Results

Table 1 presents the average percentages of responses to the item from each part of the critical set (the average percentage of "don't know" responses was 6.7%). The distribution shows a pattern of predominance of post-target intrusions, a result previously found in the "Report the letter in this color" version of the task (Botella et al., 2001). As a consequence, the average number of trials on which the rest of the calculations are based, was 57, 231, and 86, for the pre-target, correct, and post-target responses, respectively.

In order to test our main prediction, for each participant we calculated the percentage of correct T2 detections conditioned on the type of response to T1 (correct, pre-target or post-target intrusion) for each T1-T2 lag. These data were submitted to a 7x3 (lag by T1 response) ANOVA. There was a statistically significant main effect of lag, $F(6, 84) = 12.58, p < .001$, and a marginally significant main effect of response type $F(2, 28) = 3.17, p = .058$. More importantly, however, is the presence of a significant interaction between lag and response type, $F(12, 168) = 2.66, p = .003$. Figure 1 shows the nature of the interaction. As expected, the time course of the attentional blink is different when participants make a correct response to T1 in comparison to pre- and post-target intrusions, whereas these two last types of errors for T1 produce the same time course for the attentional blink. In fact, when repeating the ANOVA with only the pre- and post-target conditions, the interaction is not significant, $F(6, 84) = 0.606, p = .269$.

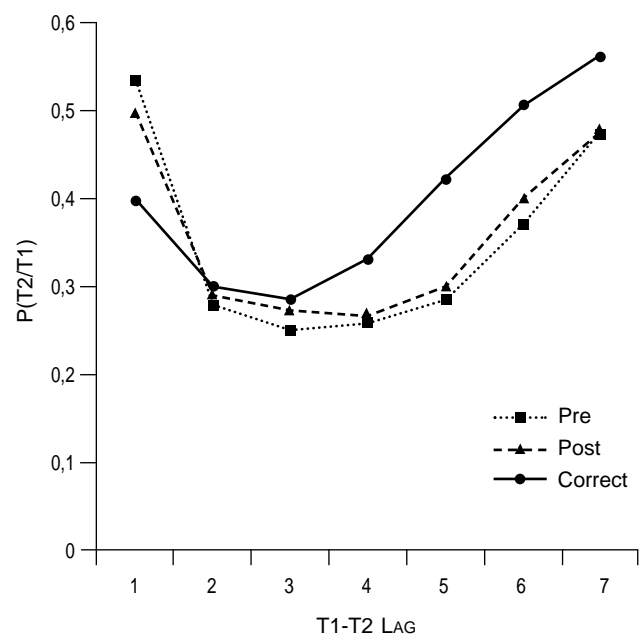


Figure 1. Mean percentages of correct detections of T2 conditionalized to the response to T1 (correct responses, pre-target intrusions or post-target intrusions) as a function of T1-T2 lag.

In order to check whether performance in lag 1 is different as a function of the type of response to T1 (different size of lag 1 sparing) we submitted those data to a one-way within-subjects ANOVA, which showed a significant effect, $F(2, 28) = 4.532, p = .02$; post hoc comparisons showed significant differences between hits and both types of intrusions ($p = .02$ in both cases) but not between the two types of intrusions ($p = .45$).

We also obtained separate cubic functions for each T1 response condition. In all three the fit was very good (R^2 : .968, .996 and .964 respectively for pre-, correct, and post-target conditions).

In order to rule out any explanation of the observed effect based on bias shifts in detecting T2 across the conditions, the false alarm rates were also analyzed. The average rates were 18.6%, 17.1% and 19.3% for pre-, hits, and post-target responses respectively; the differences were not statistically significant, $F(2, 34) = 0.273, p = .762$.

Discussion

The present study was designed to test a prediction derived from the model proposed by Botella et al. (2001) to account for what they called illusory conjunctions in the time domain. According to the model, participants can make two sequential attempts to generate a target identification response for a single target embedded in an RSVP stream of distracters (the second one is activated only when the first one fails). The first attempt is produced by focal attention and always produces correct responses when attention is focused on the target. The second attempt is generated by a sophisticated guessing mechanism that samples the levels of activation of the response features of the target, and the stimuli around it, in the series. It can produce correct responses or the type of erroneous combinations called illusory conjunctions. According to the model, the classification of trials in the correct/error dichotomy covariates with the efficiency of focal attention. We are then able to make the prediction that the interference (AB) produced upon detecting a second target will show a different time course that has a faster recovery when the response to the first target is generated from focal attention than when it is based on sophisticated guessing.

Our results match this prediction. The AB shows a longer recovery in trials where the T2 detection was conditioned to an intrusion error on T1, than when it was conditioned to a correct response. This is reflected in the significant interaction between the T1-T2 lag and the type of response to T1. The time-course of the attentional blink is different in that: (a) lag 1 sparing (Chun & Potter, 1995; Visser, Bischof, & Di Lollo, 1999) is larger following errors (performance is higher in the shortest lag when T1 intrusions are recorded); b) performance recovers earlier following correct responses. In short, assuming Botella et al's model (2001) is correct, when focal attention succeeds (most probably, when it works

efficiently), there is a smaller amount of lag 1 sparing but a speedier recovery from the AB than when the sophisticated guessing mechanism is employed to produce the response. Trials solved by the attentional focusing mechanism hypothesized in Botella et al.'s model require less time to produce a response, exactly as is predicted by the model.

Even more important is the fact that the time course of interference produced on T2 in trials with pre- and post-target intrusions to T1 are indistinguishable. According to the model, all intrusions have their origin in the same route: the sophisticated guessing mechanism triggered after focal attention fails. Therefore, the time course of the interference over the second target should be the same for pre-target and for post-target intrusions. Confirmation of this result provides specific evidence for the general structure of the model, as it accounts for the idea that hits and intrusions are typically generated by two different mechanisms. Also, the results indicate that processes resulting in pre- and post-target intrusions are probably similar in some respects, as they generated the same type and amount of interference on T2 processing. This result is very different from the predictions derived from alternative explanations of illusory conjunctions (Lawrence, 1971; McLean, Broadbent & Broadbent, 1983). According to these explanations, errors are produced because focal attention is applied to the wrong item. However, in that case, pre-target errors should show a faster recovery from AB than correct responses, whereas post-target errors should show a slower AB recovery.

In short, our results support the predictions made from Botella et al.'s (2001) model; specifically, the general structure based on two sequential attempts to generate a response. As reflected in the differential time course of the AB on a second target, success of focal attention in identifying the first target results in quick recovery from T1 processing and shorter AB. If focal attention fails, however, the sophisticated guessing mechanism takes more time to pick a response, with the results that the AB effect is lengthened in time, although it does not matter whether the eventual response is pre- or post-target intrusions.

It is worthy to highlight that the present results concur with the view of the AB being a result of a central processing bottleneck (e.g., Jolicoeur, Dell'Acqua, & Crebolder, 2001). Jolicoeur (1999) found a modulation of AB as a function of the duration of T1 processing, by manipulating the number of alternatives in a speeded choice reaction task. We have shown that the central processing stage responsible for the bottleneck is also involved in processes that Botella et al (2001) related to focal attention.

References

- Botella, J., Barriopedro, M. I., & Suero, M. (2001). A model of the formation of illusory conjunctions in the time domain. *Journal of Experimental Psychology: Human Perception and Performance*, 27, 1452-1467.

- Botella, J., & Eriksen, C. W. (1992). Filtering versus parallel processing in RSVP tasks. *Perception & Psychophysics*, *51*, 334-343.
- Chun, M. M. (1997). Temporal binding errors are redistributed in the attentional blink. *Perception & Psychophysics*, *59*, 1191-1199.
- Chun, M. M., & Potter, M. C. (1995). A two-stage model for multiple detection in RSVP. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 109-127.
- Isaak, M. I., Shapiro, K. L., & Martin, J. (1999). The attentional blink reflects retrieval competition among multiple RSVP items: Tests of the interference model. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 1774-1792.
- Jolicoeur, P. (1999). Concurrent response-selection demands modulate the Attentional blink. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 1097-1113.
- Jolicoeur, P., & Dell'Acqua, R. (1998). The demonstration of short term consolidation. *Cognitive Psychology*, *32*, 138-202.
- Jolicoeur, P., Dell'Acqua, R., & Crebolder, J. M. (2001). The attentional blink bottleneck. In K. Shapiro (Ed.), *The limits of attention: Temporal constraints in human information processing* (pp. 82-99). London: Oxford University Press.
- Lawrence, D. H. (1971). Two studies of visual search for word targets with controlled rates of presentation. *Perception & Psychophysics*, *10*, 85-59.
- McLean, J. P., Broadbent, D. E., & Broadbent, M. H. P. (1983). Combining attributes in rapid serial visual presentation tasks. *Quarterly Journal of Experimental Psychology*, *35A*, 171-186.
- Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in a RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 849-860.
- Schneider, W. (1988). Micro Experimental Laboratory: An integrated system for IBM PC compatibles. *Behavior Research Methods, Instrumentation & Computers*, *20*, 206-217.
- Seiffert, A. E., & Di Lollo, V. (1997). Low-level masking in the attentional blink. *Journal of Experimental psychology: Human Perception and Performance*, *23*, 1061-1073.
- Shapiro, K. (2001). Temporal methods for studying attention: How did we get here and where are we going? In K. Shapiro (Ed.), *The limits of attention* (pp. 1-19). New York: Oxford University Press.
- Shapiro, K. L., Arnell, K. M., & Raymond, J. E. (1997). The attentional blink. *Trends in Cognitive Sciences*, *1*, 291-296.
- Shapiro, K., Raymond, J. E., & Arnell, K. (1994). Attention to visual pattern information produces the attentional blink in RSVP. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 357-371.
- Visser, T. A. W., Bischof, W. F., & Di Lollo, V. (1999). Attentional switching in spatial and non-spatial domains: Evidence from the attentional blink. *Psychological Bulletin*, *125*, 458-469.
- Ward, R., Duncan, J., & Shapiro, K. L. (1996). The slow time-course of visual attention. *Cognitive Psychology*, *30*, 79-109.

Received July 16, 2003

Revision received March 17, 2004

Accepted March 29, 2004